

Figure 1

An amino acid sequence of human, type I, IMPDH protein

IMPDH = INOSINE-5'-MONOPHOSPHATE DEHYDROGENASE 1
514 residues

MADYLISGGT GYVPEDGLTA QQLFASADDL TYNDFLILPG FIDFIADEVD
LTSALTRKIT LKTPLISSPM DTVTEADMAI AMALMGGIGF IHHNCTPEFQ
ANEVRKVKNF EQGFITDPVV LSPSHTVGDV LEAKMRHGFS GIPITETGTM
GSKLVGVIVTS RDIDFLAEKD HTLLLSEVMT PRIELVVAPA GVTLKEANEI
LQRSKKKGKLP IVNDCDELVA IIARTDLKKN RDYPLASKDS QKQLLCGAAC
GTREDDKYRL DLLTQAGVDV IVLDSSQGNS VYQIAMVHYI KQKYPHLQVI
GGNVVTAAQA KNLIDAGVDG LRVGMGCGSI CITQEVMACG RPQGTAVYKV
AEYARRFGVP IIADGGIQTG GHVVVKALALG ASTVMMGSLL AATTEAPGEY
FFSDGVRLKK YRGMGSLDAM EKSSSSQKRY FSEGDKVKIA QGVSGSIQDK
GSIQKFVPPYL IAGIQHGCQD IGARSLSVLR SMMYSGELKF EKRTMSAQIE
GGVHGLHSYE KRLY

Figure 2

An amino acid sequence of human, type II IMPDH protein

IMPDH = INOSINE-5'-MONOPHOSPHATE DEHYDROGENASE 2
, 514 residues

The underlined region correlates with the subdomain region

MADYLISGGT SYVPDDGLTA QQLFNCGDGL TYNDFLILPG YIDFTADQVD
LTSALTKKIT LKTPLVSSPM DTVTEAGMAI AMALTGGIGF IHHNCTPEFQ
ANEVRKVKKY EQGFITDPVV LSPKDRVRDV FEAKARHGFC GIPITDTGRM
GSRLVGISS RDIDFLKEEE HDCFLEEIMT KREDLVVAPA GITLKEANEI
LQRSKKGKLP IVNEDDELVA IIARTDLKKN RDYPLASKDA KKQLLCGAAI
GTHEDDKYRL DLLAQAGVDV VVLDSSQGNS IFQINMIKYI KDKYPNLQVI
GGNVVTAAQA KNLIDAGVDA LRVGMGSGSI CITQEVLACG RPQATAVYKV
SEYARRFGVP VIADGGIQNV GHIAKALALG ASTVMMGSLL AATTEAPGEY
FFSDGIRLKK YRGMGSLDAM DKHLSSQNRY FSEADKIKVA QGVSGAVQDK
GSIHKFVPYL IAGIQHSCQD IGAKSLTQVR AMMYSGELKF EKRTSSAQVE
GGVHSLHSYE KRLF

FIGURE 3

The Subdomain of Wild-Type, Human, and Type II MPDH is Replaced with an Oligo-Peptide

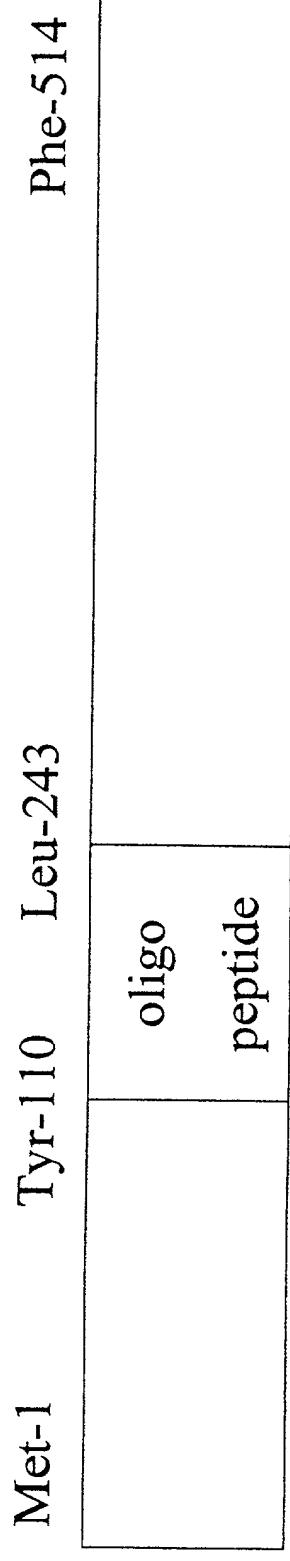


Figure 4

The amino acid sequence of the modified IMPDH-DKT polypeptide 384 residues.

The substitute tri-peptide DKT sequence is highlighted in bold print

MADYLISGGT SYVPDDGLTA QQLFNCGDGL TYNDFLILPG YIDFTADQVD
LTSALT~~K~~KKIT LKTPLVSSPM DTVTEAGMAI AMALTGGIGF IHHNCTPEFQ
ANEVRKVKKY **DKT**LLCGAAI GTHEDDKYRL DLLAQAGVDV VVLDSSQGNS
IFQINMIKYI KDKYPNLQVI GGNVVTAQA KNLIDAGVDA LRVGMGSGSI
CITQEVLACG RPQATAVYKV SEYARRFGVP VIADGGIQNV GHIAKALALG
ASTVMMGSLL AATTEAPGEY FFSDGIRLK YRGMGSLDAM DKHLSSQNR
FSEADKIKVA QGVSGAVQDK GSIIHKFVPYL IAGIQHSCQD IGAKSLTQVR
AMMYSGELKF EKRTSSAQVE GGVHSLHSYE KRLF

038536918 0631004

Figure 5

The nucleotide sequence of type II, IMPDH-DKT cDNA

atggccgactacctgattagtggggcacgtcctacgtgccagacgacggactcacagcacagcagctct
tcaactgcggagacgcctcacctacaatgacttctcatttcctggcacatcgacttcactgcaga
ccaggtggacctgacttctgctgaccaagaaaatactctaagaccccactggttcctccatq
gacacagtacacagaggctggatggccatagcaatggcgttacaggcggattggcttcatccaccaca
actgtacacacctgaattccaggccaatgaagttcgaaagtgaagaaatatgacaagacccctgctgttgg
ggcagccattggcactcatgaggatgacaagtataaggctggacttgctcgccaggctggatgt
gtgggtttggactctccagggaaattccatcttccagatcaaatatgatcaagtacatcaaagacaat
accctaatctcoaagtcatggaggcaatgtggtaactgctgcccaggccaagaacctcattgatgcagg
tgtggatgccctgcgggtggcatggaaagtggctccatctgcattacgcaggaagtgtggcctgtgg
ggcccccaagcaacagcagtgtacaagggtcagagtatgcacggcgttgggttccggcattgctg
atggaggaatccaaaatgtgggtcatattgcgaaagcctggccctggggcctccacagtcatgatgg
ctctctccctggctgcaccactgaggcccctggtaataacttctttccatgggatccggctaaagaaa
tatcgccgtatgggtctctcgatgccatggacaagcacctcagcagccagaacagatattcagtgaag
ctgacaaaatcaaagtggccaggagtgctggctgtcagagacaagggtcaatccacaaatttgc
cccttacctgattgtggcatccaacactcatgccaggacattgggtgccaagagcttgaccgaatccga
ccatgatgtactctgggagcttaagttgagaagagaacgtcctcagcccagggtggaaggtggcgtcc
atagcctccattcgtatgagaagcggctttctga

Figure 6

The amino acid sequence of the modified IMPDH-SPS
polypeptide 384 residues.

The substitute tri-peptide SPT sequence is highlighted in
bold print.

MADYLISGGT SYVPDDGLTA QQLFNCGDGL TYNDFLILPG YIDFTADQVD
LTSALTKKIT LKTPLVSSPM DTVTEAGMAI AMALTGGIGF IHHNCTPEFQ
ANEVRKVKKY **SPS**LLCGAAI GTHEDDKYRL DLLAQAGVDV VVLDSSQGNS
IFQINMIKYI KDKYPNLQVI GGNVVTAQA KNLIDAGVDA LRVGMGSGSI
CITQEVLACG RPQATAVYKV SEYARRFGVP VIADGGIQNV GHIAKALALG
ASTVMMGSLL AATTEAPGEY FFSDGIRLKK YRGMGSLDAM DKHLSSQNRY
FSEADKIKVA QGVSGAVQDK GSIIHKFVPYL IAGIQHSCQD IGAKSLTQVR
AMMYSGELKF EKRTSSAQVE GGVHSLHSYE KRLF

09853648 004004

FIGURE 7

The nucleotide sequence of type II, IMPDH-SPS cDNA

atggccgactacctgatttagtggggcacgtcctacgtgccagacgacgactcacagcacagcagctct
tcaactcgagacggcctcacataatgactttctcattccctgggtacatcgacttcactgcaga
ccagggtggacctgacttctgctgaccaagaaaatcactctaagaccactgggttcctccatg
gacacagtacagaggctggatggccatagcaatggcgcttacaggcggtattggctcatecaccaca
actgtacacctgaattccaggccaatgaagtcggaaagtgaagaaatattctccgagcctgtgtgg
ggcagccattggcactcatgaggatgacaagtataggctggacttgctgcccaggctgtgtggatgta
gtgggtttggacttcccaggaaattccatcttccagatcaaatatgatcaagtacatcaaagacaat
accctaatctcoaagtcatggaggcaatgtggactgtctgcccaggccaagaacaccttattgatgcagg
tgtggatgccctgcgggtggcatggaaagtggctccatctgcattacgcaggaaagtgtggctgtgg
cgcccccaagcaacagcagtgtacaagggtcagagtatgcacggcgcttgggttccggcattgctg
atggaggaaatccaaaatgtggatcatattgcgaaagccttggcccttgggcctccacagtcatgatgg
ctctctcctggctgccaccactgaggccctggtaataacttcttccatggatccggctaaagaaa
tatcgccgtatgggtctcgatgccatggacaagcacctcagcagccagaacagatatttcagtgaag
ctgacaaaaatcaaagtggcccaggaggtgtctgggtgtgcaggacaaagggtcaatccacaaatttgc
cccttacctgattgtggcatccaaacactcatgccaggacatggtgccaaagagactgaccacaacttgc
ccatgatgtacttggggagcttaagttgagaagagaacgtcctcagccagggtggaaagggtggcgtcc
atagcctccattcgatgagaagcggtttctga

卷之三

Figure 8

The amino acid sequence of the type II, modified IMPDH-GSG polypeptide

The substitute tri-peptide GSG sequence is highlighted in bold print

MADYLISGGT SYVPDDGLTA QQLFNCGDGL TYNDFLILPG YIDFTADQVD
LTSALTKKIT LKTPLVSSPM DTVTEAGMAI AMALTGGIGF IHHNCTPEFQ
ANEVRKVKKY **GSG**LLCGAAI GTHEDDKYRL DLLAQAGVDV VVLDSSQGNS
IFQINMIKYI KDKYPNLQVI GGNVVTAQA KNLIDAGVDA LRVGMGSGSI
CITQEVLACG RPQATAVYKV SEYARRFGVP VIADGGIQNV GHIAKALALG
ASTVMMGSLL AATTEAPGEY FFSDGIRLKK YRGMGSLDAM DKHLSSQNR
FSEADKIKVA QGVSGAVQDK GSIHKFVPYL IAGIQHSCQD IGAKSLTQVR
AMMYSGELKF EKRTSSAQVE GGVHSLHSYE KRLF

098535318 • 0564004

Figure 9

The nucleotide sequence of type II, IMPDH-GSG cDNA

atggccgactacctgattagtggggcacgtcctacgtgccagacgcacggactcacagcacagcagctctcaactgcggagacggcctcacctacaatgacttctcattctccctgggtacatcgacttcactgcaga
ccagggtggacacctgactctgctgaccaagaaaatcactctaagaccccactggttccctccatg
gacacagtacacagaggctggatggccatagcaatggcgctiacaggcgtattggcttcatccaccaca
actgtacacacctgaattccaggccaatgaagttcgaaagtgaagaaatatggttccggcctgtgtgg
ggcagccattggcactcatgagatgacaagtataaggctggacttgctcgccaggctgtgtggatgt
gtgggtttggactcttccagggaaattccatcttccagatcaaatatgtcaagtgatcaatcaaagacaat
accctaattctccaagtcatggaggcaatgtggtaactgctgcccaggcaagaacaccttattgatgcagg
tgtggatggccctgcgggtggcatggaaagtggctcatctgcattacgcaggaagtgcgtggcctgtgg
cgcccccaagcaacagcagtgtacaagggtcagagtatgcacgcgccttgggtgttccggcattgt
atggaggaatccaaaatgtgggtcatattgcgaaaggcctggccctgggcctccacagtcatgtatgg
ctctctctggctgcaccactgaggcccctggtaataacttctttccatggatccggctaaagaaaa
tatcgcggtatgggtctctcgatggacaaggcacctcagcagccagaacagatattcagtgt
ctgacaaaatcaaagtggcccaggagtgctgggtgcaggacaaggtaatccacaaaattgt
cccttacctgattgtggcatccaacactcatgccaggacattgtgccaagagcttgaccacaaagtcga
gccatgatgtactctggggagcttaagttgagaagagaacgtcctcagcccagggtggaggtggcgtcc
atagcctccattcgatgagaagcggctttctga

Figure 10

The amino acid sequence of the modified IMPDH-SPT polypeptide 384 residues.

The substitute tri-peptide SPT sequence is highlighted in bold print.

MADYLISGGT SYVPDDGLTA QQLFNCGDGL TYNDFLILPG YIDFTADQVD
LTSALTKKIT LKTPLVSSPM DTVTEAGMAI AMALTGGIGF IHHNCTPEFQ
ANEVRVKKKY **SPT**LLCGAAI GTHEDDKYRL DLLAQAGVDV VVLDSSQGNS
IFQINMIKYI KDKYPNLQVI GGNVVTAQA KNLIDAGVDA LRVGMGSGSI
CITQEVLACG RPQATAVYKV SEYARRFGVP VIADGGIQNV GHIAKALALG
ASTVMMGSLL AATTEAPEGEY FFSDGIRLKK YRGMGSLDAM DKHLSSQNRY
FSEADKIKVA QGVSGAVQDK GSIIHKFVPYL IAGIQHSCQD IGAKSLTQVR
AMMYSGELKF EKRTSSAQVE GGVHSLHSYE KRLF

Figure 11

The nucleotide sequence of type II, IMPDH-SPT cDNA

卷之三

Figure 12

The nucleotide sequence of type II, IMPDH-SPTQ cDNA

atggccgactacctgatttagtggggcacgtcctacgtgccagacgacgactcacagcacagcagcttc
tcaactgcggagacggcctcacataatgactttctcattccctgggtacatcgacttcactgcaga
ccagggtggacctgacttctgtctgaccaagaaaatcaactctaagaccactgggttcctccatg
gacacagtacacagaggctggatggccatagcaatggcgcttacaggcggtattgggttcatcaccaca
actgtacacactgaattccaggccaatgaagttcgaaagtqaagaaatattctccgactcagctgt
tggggcagccattggactcatgaggatgacaagtgataggctggacttgcgcggcaggctgggt
gtagtgggtttggactcttccaggaaattccatctccagatcaatatgatcaagtgatcaaagaca
aataccctaatctccaagtcatggaggcaatgttgtcactgtgcctccaggcaagaacactcattgt
aggtgtggatgcctgcgggtggcatggaaagtggctccatctgcattacgcaggaagtgtggctgt
ggcgcccccaagcaacagcagtgtacaaggtgtcagagtatgcacggcgcttgggtgttccggtatt
ctgatggaggaatccaaaatgtgggtcatattgcgaaagcctggccctgggcctccacagtcatt
ggctctccctggctgccaccactgaggcccctgttgaataacttcttccatggatccgctaaag
aaatatcgccgtatgggtctcgatgccatggacaagcacctcagcagccagaacagatattcagtg
aagctgacaaaatcaaagtggccaggaggtgtctgggtgtcaggacaagggtcaatccacaaatt
tgtcccttacctgattgtgcattccaaactcatgccaggacattgggtgccaagagacttgacccaagtc
cgagccatgtgtactctggggagcttaagttgagaagagaacgtcctcagcccagggtggaaagggtggcg
tccatagccattcgtatgagaagcgctttctga

Figure 13

The amino acid sequence of the modified type II, IMPDH-AGRP
polypeptide
385 residues

The substitute tetra-peptide AGRP sequence is highlighted
in bold print.

MADYLISGGT SYVPDDGLTA QQLFNCGDGL TYNDFLILPG YIDFTADQVD
LTSALTKKIT LKTPLVSSPM DTVTEAGMAI AMALTGGIGF IHHNCTPEFQ
ANEVRVKVKY **AGRP**LLCGAA IGTCHEDDKYR LDLLAQAGVD VVVLDSQGN
SIFQINMIKY IKDKYPNLQV IGGNVVTAAQ AKNLIDAGVD ALRVGMGSGS
ICITQEVLAC GRPQATAVYK VSEYARRFGV PVIADGGIQN VGHIAKALAL
GASTVMMGSL LAATTEAPGE YFFSDGIRLK KYRGMGSLDA MDKHLSSQNR
YFSEADKIKV AQGVSGAVQD KGSIHKFVPY LIAGIQHSCQ DIGAKSLTQV
RAMMYSGEELK FEKRTSSAQV EGGVHSLHSY EKRLF

Figure 14

The nucleotide sequence of type II, IMPDH-AGRP

atggccgactacctgatttagtggggcacgtcctacgtgccagacgacgactcacagcacagcagctctcaactgcggagacggcctcacataatgactttctcatttcctggtatcatcgacttcactgcaga
ccagggtggacactgacttctgctgaccaagaaaatcactctaagaccccactggttcctccatgacacagtacacaggctggatggccatagcaatggcgcttacaggcgtattggctcatcaccaca
actgtacacactgaattccaggccaatgaagttcgaaagtgaagaaatatgctggcgtccgctgttgatggggcagccattggcactcatgaggatgacaagtataggctggacttgcgcggctggatggat
gttagtggtttggacttccaggaaattccatttccagatcaaatatgatcaagtagatcaaagaca
aataccctaatctccaagtcatggaggcaatgtggtaactgctgcccaggccaagaacacctattgatgc
aggtgtgatgcccctgccccatgggaagtggctccatctgcattacgcaggaaagtgtggcctgt
ggccggcccaagcaacagcagtgtacaagggtcagagtatgcacggcgttgggttccggcattg
ctgatggaggaatccaaaatgtggctatattgcggaaagcctggccctgggcctccacagtcattgat
ggcctctccctggctgccaccactgaggcccctggtaataacttcttcgcattggatccggctaaag
aaatatcgcggatgggtctcgatggccatggacacactcagcaggccagaacagatattcgt
aagctgacaaaatcaaagtggcccaggaggtgtctgggtctgtcaggacaaagggtcaatccacaaatt
tgtcccttacctgattgctggcatccaaacactcatgccaggacattgggtgccaagagacttgacccaagtc
cgagccatgtgtactctggggagcttaagttgagaagagaacgtcctcagcccaaggtggaaaggtggcg
tccatagcctccatcgtatgagaagcggctttctga

卷之三

Figure 15

The amino acid sequence of type II, modified IMPDH-NSPL polypeptide

The substitute tri-peptide is highlighted in bold print

MADYLISGGT SYVPDDGLTA QQLFNCGDGL TYNDFLILPG YIDFTADQVD
LTSALTKKIT LKTPLVSSPM DTVTEAGMAI AMALTGGIGF IHHNCTPEFQ
ANEVRKVKKY **NSPL**LLCGAA IGTHEDDKYR LDLLAQAGVD VVVLDSQGN
SIFQINMIKY IKDKYPNLQV IGGNVVTAAQ AKNLIDAGVD ALRVGMGSGS
ICITQEVLAC GRPQATAVYK VSEYARRFGV PVIADGGIQN VGHIAKALAL
GASTVMMGSL LAATTEAPGE YFFSDGIRLK KYRGMGSLDA MDKHLSSQNR
YFSEADKIKV AQGVSGAVQD KGSIHKFVPY LIAGIQHSCQ DIGAKSLTQV
RAMMYSGELK FEKRTSSAQV EGGVHSLHSY EKRLF

Figure 16

The nucleotide sequence of type II, IMPDH-NSPL cDNA

atggccgactacctgatttagtggggcacgtcctacgtgccagacgacgactcacagcacagcagctct
tcaactgcggagacggcctcacataatgactttctcattccctggtacatcgacttcactgcaga
ccaggtggacctgacttctgctgaccaaaatcactctaagaccccactggttcctccatg
gacacagtacacaggctggatggccatagcaatggcgcttacaggcgtattggcttcatccaccaca
actgtacacctgaattccaggccaatgaagtccggaaagtgaagaaatataactctccgcttctgt
tggggcagccattggactcatgaggatgacaagtataggctggacttgcgcggcaggctggat
gtatgtgtttggactcttccaggaaattccatcttccagatcaatatgatcaagtacatcaaagaca
aataccctaatacttccaagtcatggaggcaatgtggtaactgctgccaggcaagaacacttattgtgc
aggtgtggatgcctgcgggtggcatggaaagtggctccatctgcattacgcaggaaagtgcgt
ggcgcccccaagcaacagcagtgtacaagggtcagagtatgcacggcgcttgggttccggtcattg
ctgatggaggaatccaaaatgtgggtcatattgcggaaagcctggccctgggcctccacagtc
ggctctccctggctgccaccactgaggccctggtaataactcttccatggatccgctaaag
aaatatcgccgtatgggtctcgatgccatggacaagcacctcagcagccagaacagatattc
aagctgacaaaatcaaagtggccaggagtgctgtgctgtgcaggacaagggtcaatccacaatt
tgtcccttacctgattgtggcatccaaacactcatgccaggacattgggtccaaagagacttgc
cggccatgtgtactctggggagcttaagttgagaagagaacgtcctcagcccagggtggaaagg
tggcgccatgcctccatcgatgagaagcggctttctga

Figure 17

The amino acid sequence of the type I modified IMPDH-DKT polypeptide

The substitute tri-peptide DKT is highlighted in bold print

MADYLISGGT GYVPEDGLTA QQLFASADGL TYNDFLILPG FIDFIADEVD
LTSALTRKIT LKTPLISSPM DTVTEADMAI AMALMGGIGF IHHNCTPEFQ
ANEVRKVKKF **DKT**LLCGAAV GTREDDKYRL DLLTQAGVDV IVLDSSQGNS
VYQIAMVHYI KQKYPHLQVI GGNVVTAQA KNLIDAGVDG LRVGMGCGSI
CITQEVMACG RPQGTAVYKV AEYARRFGVP IIADGGIQTG GHVVKALALG
ASTVMMGSLL AATTEAPGEY FFSDGVRLKK YRGMGSLDAM EKSSSSQKRY
FSEGDKVKIA QGVSGSIQDK GSIQKFVPYL IAGIQHGCQD IGARSLSVLR
SMMYSGELKF EKRTMSAQIE GGVHGLHSYE KRLY

Figure 18

A schematic representation of the distance that the substitute oligo-peptides are designed to span in a folded modified IMPDH polypeptide.

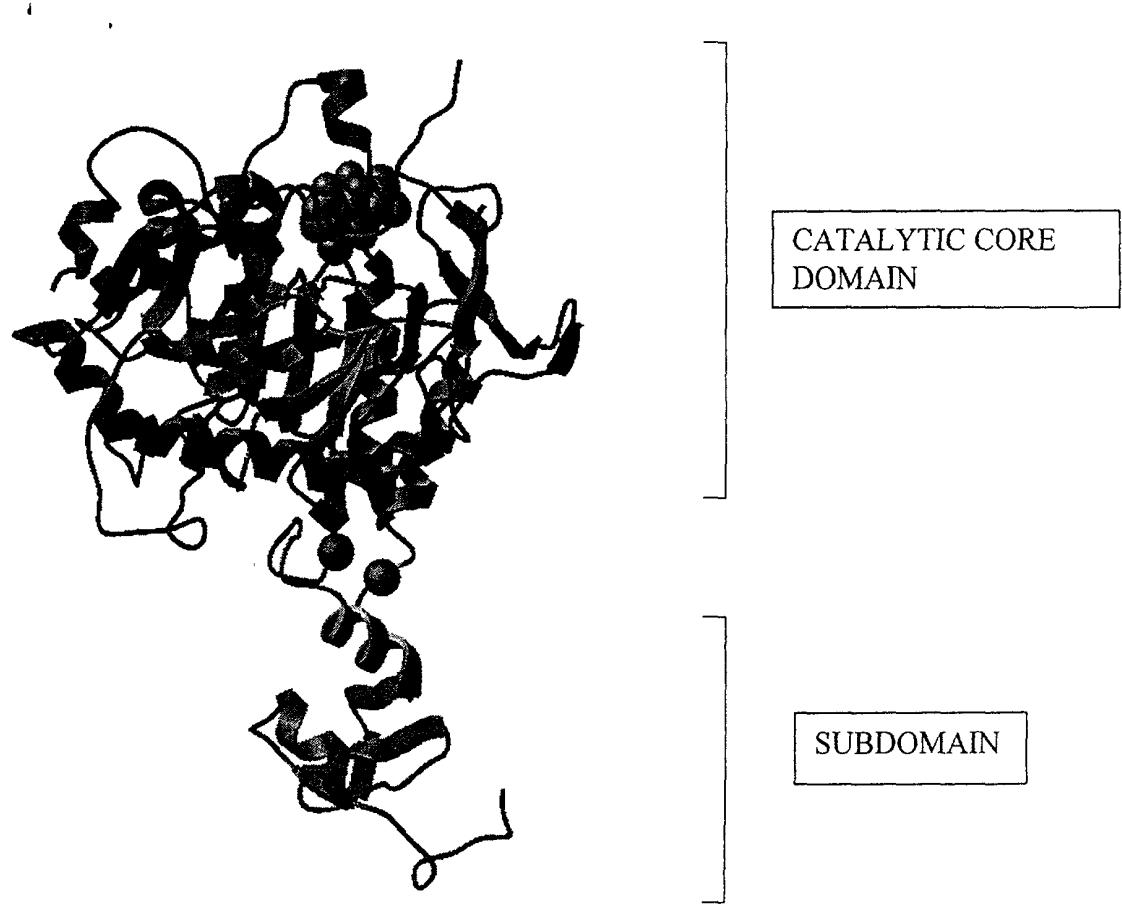


Figure 19

The Rate of NADH Production at 37 °C for Wild-Type IMPDH (type II) and Various Modified IMPDH Multimers.

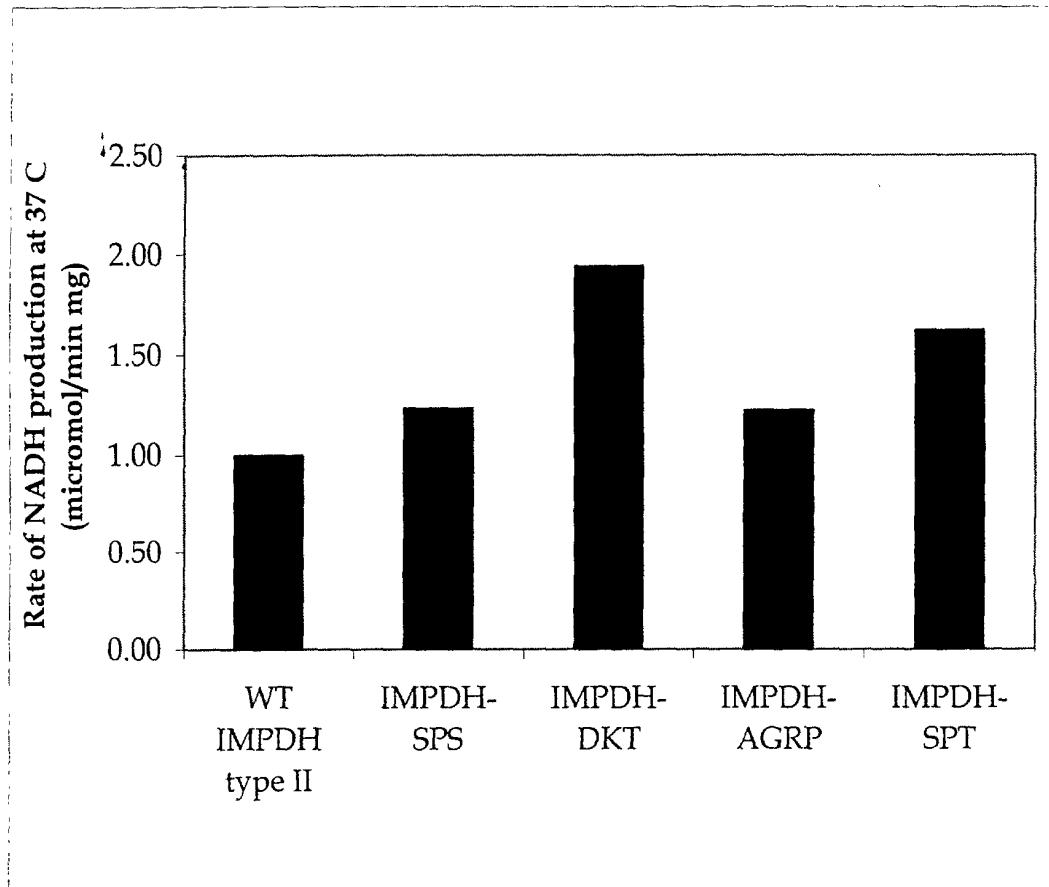


Figure 20

MPA Inhibits the Activity of Various Modified IMPDH Polypeptides.

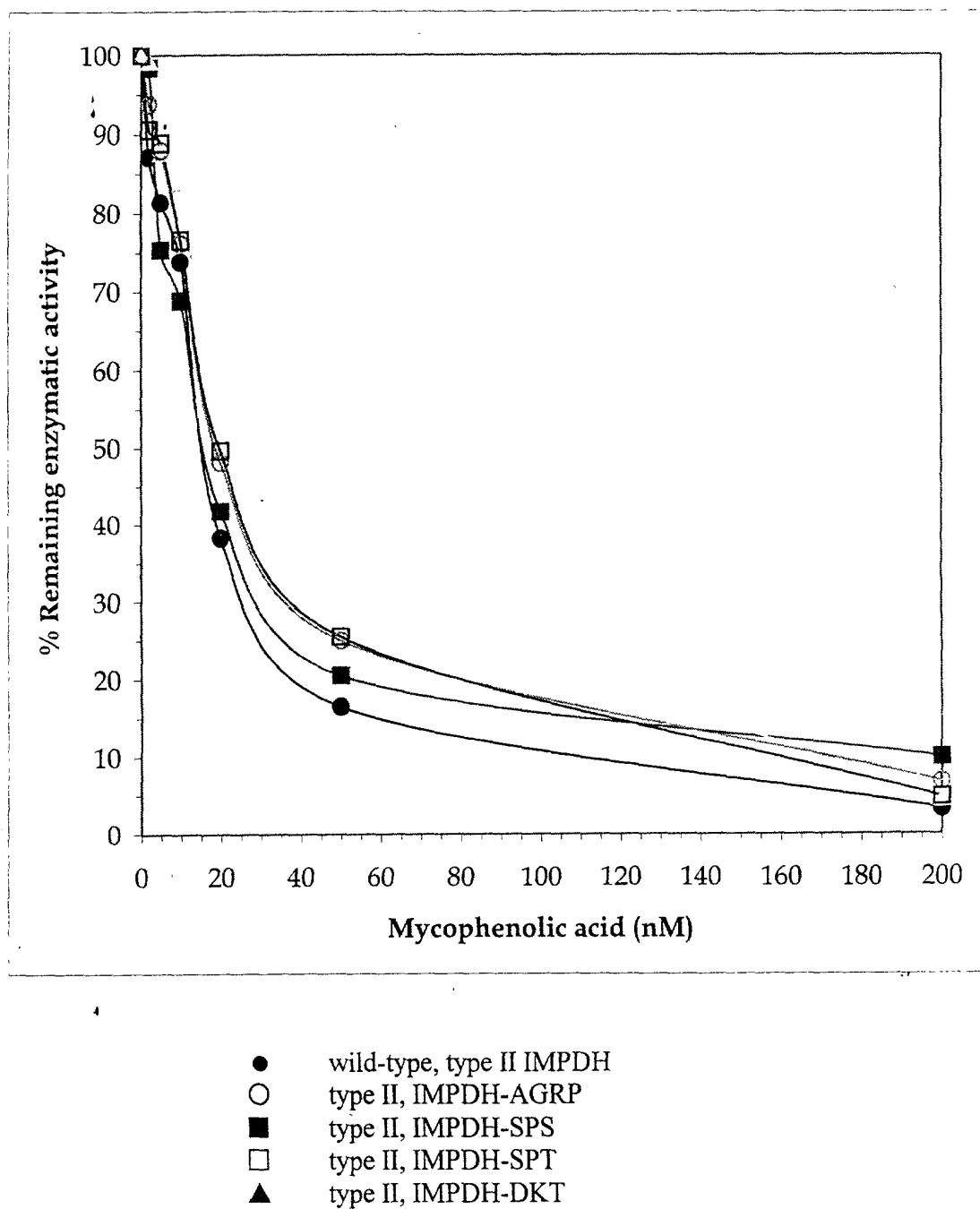
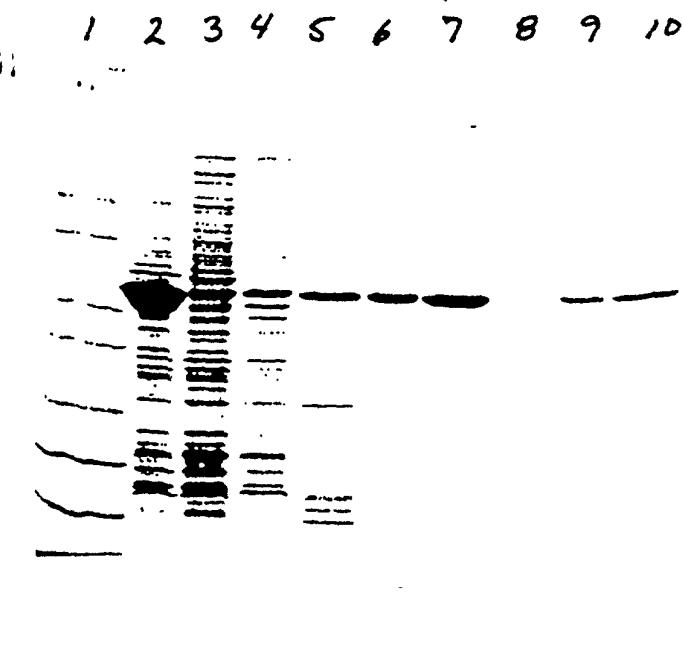


FIGURE 21



- | Lane # | |
|--------|--|
| 1: | Novex standards (25 μ L) |
| 2: | Total cell lysate (1 μ L, 12 μ g; before ultracentrifugation) |
| 3: | Soluble lysate (2.5 μ L, 10 μ g; after 100,000 \times g, 1 hr, 4 °C) |
| 4: | Unbound AE sample (13 μ L, ~10 μ g) |
| 5: | Blue dye column, frs. #32-70 (25 μ L, 4.5 μ g) |
| 6: | IMP affinity column, IMP eluted (10 μ L, ~1.5 μ g) |
| 7: | IMP affinity column, IMP eluted (20 μ L, ~3.0 μ g) |
| 8: | Unbound protein to IMP column (25 μ L, ~2.0 μ g) |
| 9: | IMP affinity column (from AE fr. #1-10, 10 μ L, 1.2 μ g) |
| 10: | IMP affinity column (from AE fr. #1-10, 25 μ L, 3.0 μ g) |

FIGURE 22

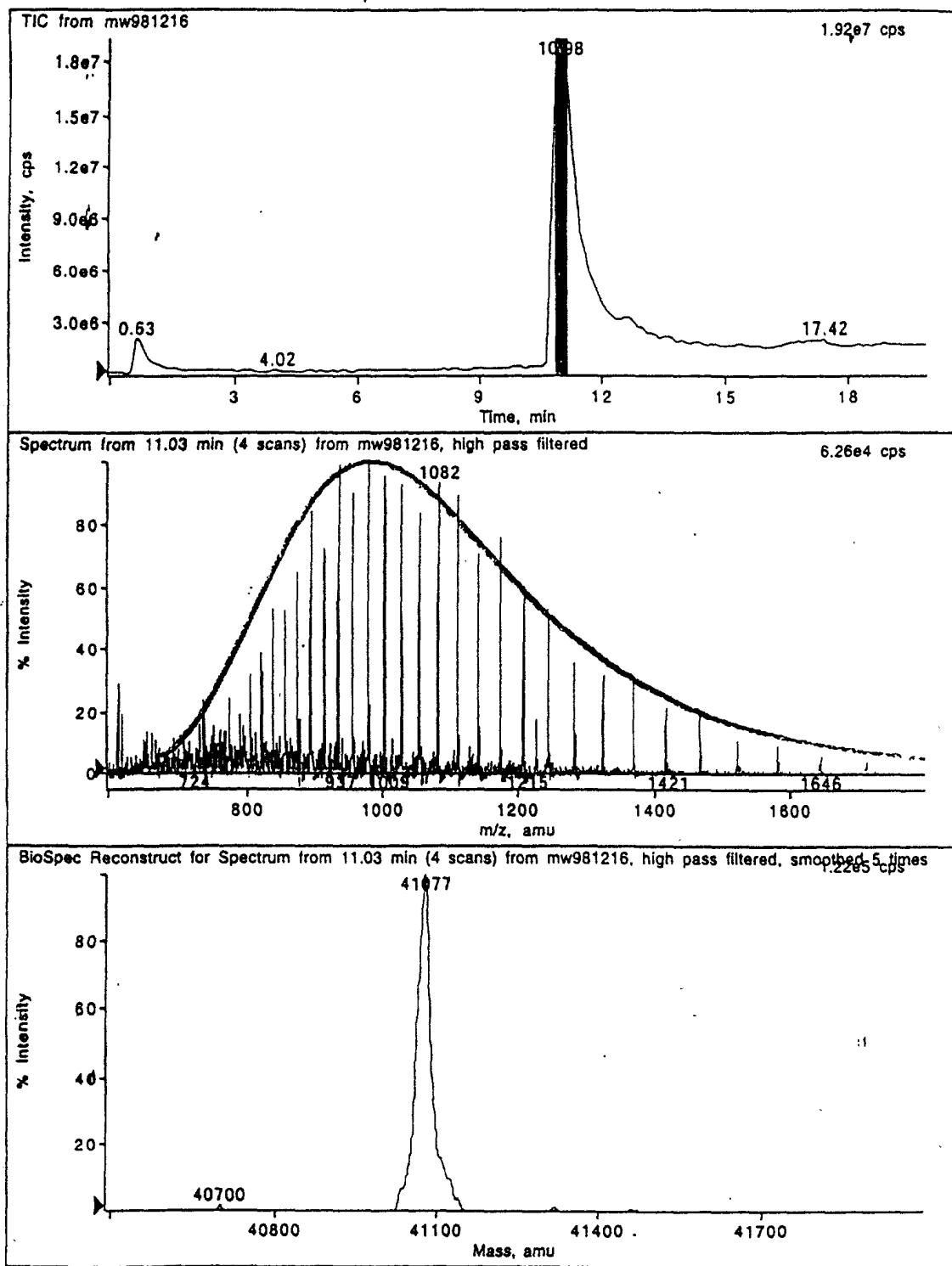


FIGURE 23

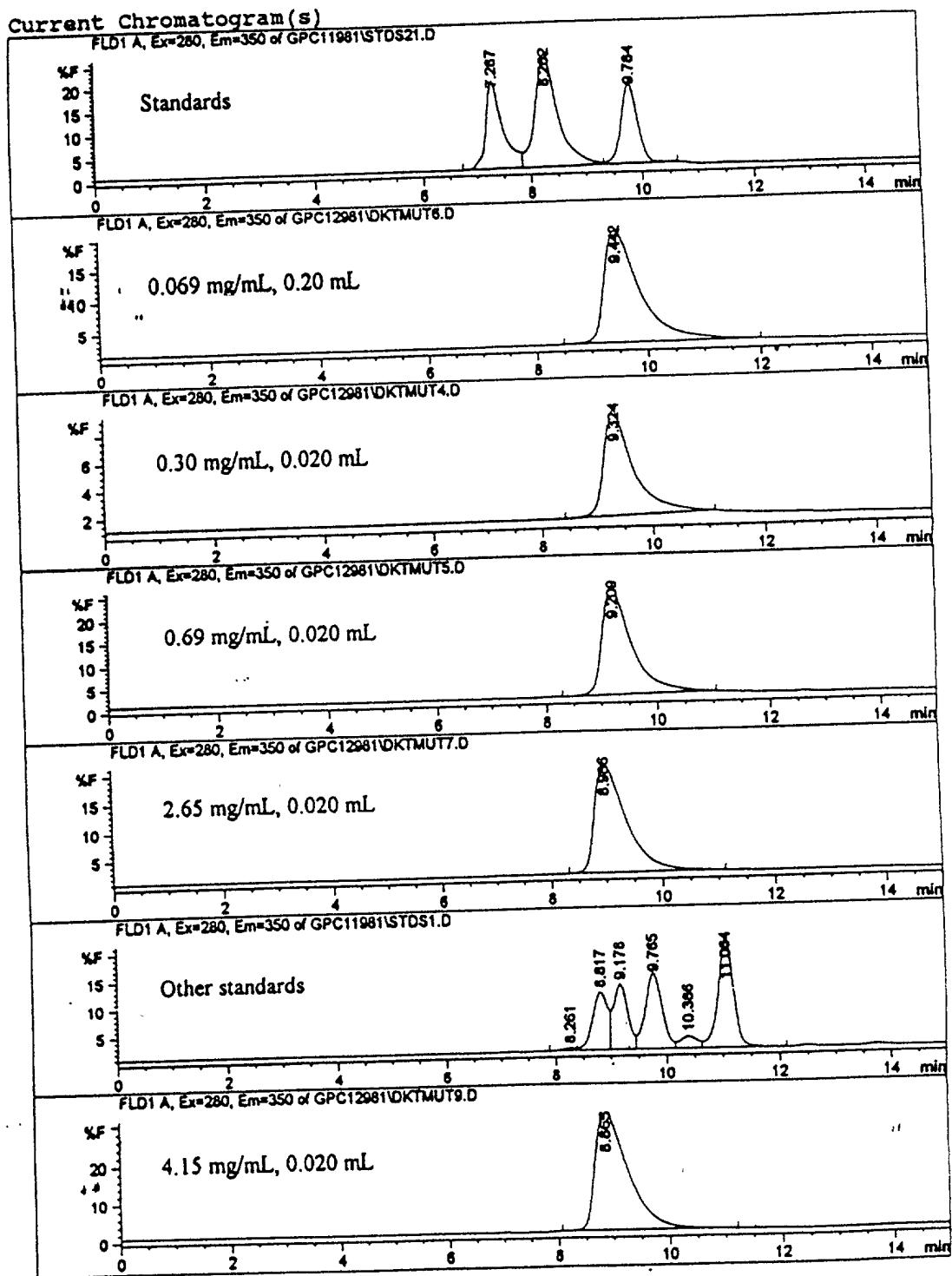


Figure 24

The nucleotide sequence of type I, IMPDH-DKT

atggcgactacctgatcagcgccggcaccggctacgtgcccaggatggctaccgcgcagcagctct
tcgcgcgcgcacggcctcacataacgacttctgattctccaggattcatagacttcatagctga
tgagggtggacctgacctcagccctgaccggaaagatcacgctgaagacgcactgatctcccccatt
gacactgtgacagaggctgacatggccatggctctgatggaggattggttcattcaccaca
actgcaccccaagatccaggccaacgagggtcgaaaggtaagaagttgacaaaaccctgtctgtgg
ggcagctgtggcaccgtgaggatgacaaataccgtctggacactgctcaccggggcgacgtc
atagtcttggactcgtnccaaggaaattcggttatcaaatacgccatggcattacatcaaacaagaat
accccacctccagggtattggggaaacgtgtgacagcagcccaggcaagaacctgattgatgtgg
tgtggacggctgcgtggcatggctcggtccatctgatcaccaggaaatggatggcctgtgg
cggccccaggcactgtgttacaagggtggctgatgcccggcgcttgggtgtgccatcatagcc
atggccgcattccaggaccgtggacacgtgtcaaggccctggcattggacgcctccacagtatgtgg
ctccctgtggccgcaactacggaggcccggcagttctcagacgggtcggtcaagaag
taccgggcattggctactggatggcatggagaagagcagcagccagaaacgataacttcagcgagg
gggataaaagtgaagatcgccgcagggtgtctcggtccatccaggacaaaggatccattcagaagttcg
gccctacctcatcgaggcatccaacacggctgcaaggatatcgccggccgcagcctgtctgtcc
tccatgtgtactcaggagagctcaagttgagaagcggaccatgtcgcccgatgggtgtgtcc
atggcctgcactttacgaaaagcggctgtactga